

SEQUENCE LISTING

SEQ ID NO:1 IND1 genomic

Sequence Range: 1 to 3856

5
10
15
20
25
30
35
40
45
50
55
60
65

CTCTAGACCA TCTACTATCC GGTGTTGAC CCTTAAAGCT TTTGAAGACT ACTAGAATAA 60
TGCAAATACC ATATGTCCAT ATCCATCCTT TTCTTTTGTT TGAAGTGAAC ATTCTAATTT 120
TGTAAGAGAA AAAACCTTAT GTTAATATCA CCGTAGGCAA AAAAAATATC TCATCATATT 180
AAATTTTAT TATAAGATTA TACATTCTCT CGTTGTAAGA GTTACTCCAA TTGCAAGTGT 240
TGTATTAAGT AATAAAAGG ACGAAAGTAG GAAGCTTATA ATTAATTGAT GTTGCATAGT 300
ACTGGTATAT TGTGATGAA TATAACAAGT ATGAACATTA ATGCATGAAA CGGGGTATTT 360
TGTCTGAAC TCATTAAAGG CAATGTGAAA AGAAGATGTG AGGTCTCATT TTGAAAATTT 420
ATCTTCTAGC TTTGTCGATT TTAATCTAT GAAATGAACG CAACATATAG AAATTTTCATG 480
TGGACAACGA CATTTAGACG GTATCTTAAT TAGACCGATT AATTAGTAAT ATACTTATAT 540
ATATAATTAG TGGTGATTAT AAGTTTACTT ATCCACTTGA GAATTTAAAC AATGGGCAAT 600
ACCTTAATGT CGAAAGAAGC CGTCCCACT TCGTGTAATG AGTTATGGGG GAGAGATCCT 660
GTTAAATCGT CAAATAAAAC AACTTAAGAA CTAGAAATTG ACACCAAAAA TCATAAGAG 720
AACGTTGAAG AAGTCATTTA TGGTATCCAG CTCATATTTT CTAGCTAAGA TCAAATCAAG 780
GCCGTTGAAA GGGCTTGTA GAAATGTGG AAGAAACCGT GGGGTTTAGA AGAAAGACAA 840
GAAATAGAAG AACAAATGATG TTAATTTGCC TATTTTGGTG TATAGGAGTT GTCAAAAGAG 900
GAGAGAGAGA AGAAAATTAG GTCAAAATAA TGAGCACTAA AAATGGAGAC ATGTGTTGAG 960
TAACTATTAC AAGAGCGACT TATGCTTCCT TATGGCAATG ATATCCAAAC CAAAGTGCAA 1020
CGCTCCTTTT TTGCCCTAAT TTCGTAAAGT CTCTCTCCTT CTTCGTCCTT AGGAAAAACC 1080
CTAGAAATTT AATCCCTTGT TCTTGATCTT GCTTTTGGAG TAACCATGAT TTTGACCACA 1140
CACTATTTCT TCTATCTTTT GTGGTCTATA GGATTTTGCT TTATATGTGT TTCTTGATTT 1200
GCTCCGTACG TACGTATACG AATTTAAATG GTTATAACAA GGTTTATATA AACTAGCACA 1260

	AATGAGTCCA	TGAAATTTGT	TAGCGAAAAA	GGTAGAAATA	TATTGAGTCT	TTAAACGGCA	1320
5							1380
	ATATATATAA	TTTTGCTGCA	AAACTTAGCT	TTAATCATGA	TCTAATGATA	TTTTCTTTAA	1440
10							1500
	TTTCCTTTGC	CAAATTAATC	ACATGCACGG	ATTTTTGGCA	AGTTATGTGT	CGAATTCCTC	1560
							1620
15							1680
	CATTCACACA	ACACTAAACT	TAATTAGAAC	TCTAGGAAAT	ATTTTAAAAAT	GACAACTTTA	1740
							1800
20							1860
	TCGAAAAAAA	TTTAGTTATG	AAAACAATTC	CAGAATTAAA	CATGAGCTAT	ATAATTTAAG	1920
							1980
25							2040
	ATAAAATGAA	GTAATATTGA	TATGTATGTA	ATAACATATC	TGATTGCGGT	AAAAAAAAC	2100
							2160
30							2220
	ATATCTGATT	AAATTGTTCA	TGCAGGCCCA	TGTCACTATG	ATGTCATCAC	GTTTTTATTT	2280
							2340
35							2400
	TCACAATAAC	TAATATATAT	TCAAAAAAAT	AGTTTTGTCA	GATTAAATTT	TTTTTGGTGG	2460
							2520
40							2580
	TCAGCTTTCT	CCAACCTACT	AAACTAGTTT	GGAATGTTCT	CTTCTTTATT	TTTCTTTTTC	2640
							2700
45							2760
	TTGATTTCTT	ATGTTTTTTA	TTTATGGAAT	TTTAAGACGG	ATTGTTTAGG	TCGTTTCTCT	2820
							2880
50							2940
	CTTTTCTTGT	TTTCTAAAGT	TACTTTTGTA	AACTCATCTC	CTCCCAATTA	GACAGTCAAT	3000
							3060
55							3120
	CATATAGTTA	TCTTTTAATA	TATGTCTAGT	TGATAAAAAA	AATGAAAAAA	TACTGGTGGT	3180
							3240
60							3300
	AGTTCTACTA	ATGTTTGTGT	AAAAAATCTG	ATATTATGAA	TCTAATCAAT	TTCTTTGATC	3360
							3420
65							3480
	GTATAATGTG	GGTTAAATTT	AGTAATTTTT	TACATAAATA	AGAACTGTAA	TGTTGATGTA	3540
							3600
70							3660
	TATTGGGGAA	TCAGTATATT	AGCTTGGGTA	ACTATACTTC	TGGAAATACT	TGAAGATTTA	3720
							3780
75							3840
	ACTATTTGCA	AAATTATAAT	TTAGTCCCGA	AAAATACAGA	CGACGGGACA	CGACAACATA	3900
							3960
80							4020
	TAAGCAGGTT	TGAATCTTGG	AAAATTTTGT	ATACATAACC	TATATAAATA	CTAATGTTCT	4080
							4140
85							4200
	GGTTGGGTTC	AAAAGCCTTT	TCAAAAGTTC	CATTTTTTTAA	ATTCAAGGAC	ATTTTACATA	4260
							4320
90							4380
	GGAAATAAGT	TGAGTCATAA	AAAATAATGG	TTATTTTGTA	AGGTTTTTTT	TTTGATTAAA	4440
							4500
95							4560
	ACGCACATAT	TAAGAAGTTA	GTTTTTTTTTC	ACTACCAAT	ATCAATTAAT	TTAAAACCAT	4620
							4680
100							4740
	GCAACCATTC	ATAAAACAAT	ACTATTAAAG	AATATAAATA	ATCACAAAAT	ATTAAATACA	4800
							4860
105							4920
	CTTAAAATTT	ACATATAAAT	TTACAAAACA	TCTAATTAAT	TGAAACAGAA	AGGAAAAGGT	4980
							5040

AAAATATATC ATAAAATGAG ACATATATCC TATAAAAAAA AAATGAGGCA TATGAAGTAA
 2700
 5 ATAATAAGAG ACATGCATGT AAGCATTTCGG TTAATTAATC GAGTCAAAGA TATATATCAG
 2760
 TAAATACATA TGTGTATATT TCTGGAAGAAA GAATATATAT ATTGAGAAAT AAGAAAAGAT
 2820
 10 GAAAATGGAA AATGGTATGT ATAAAAAGAA AGGAGTGTGC GACTCTTGTG TCTCGTCCAA
 M E N G M Y K K K G V C D S C V S S K>
 2880
 15 AAGCAGATCC AACCACAGCC CCAAAAGAAG CATGATGGAG CCTCAGCCTC ACCATCTCCT
 S R S N H S P K R S M M E P Q P H H L L>
 2940
 20 CATGGATTGG AACAAAGCTA ATGATCTTCT CACACAAGAA CACGCAGCTT TTCTCAATGA
 M D W N K A N D L L T Q E H A A F L N D>
 3000
 25 TCCTCACCAT CTCATGTTAG ATCCACCTCC CGAAACCCTA ATTCACCTGG ACGAAGACGA
 P H H L M L D P P P E T L I H L D E D E>
 3060
 30 AGAGTACGAT GAAGACATGG ATGCGATGAA GGAGATGCAG TACATGATCG CCGTCATGCA
 E Y D E D M D A M K E M Q Y M I A V M Q>
 3120
 35 GCCCGTAGAC ATCGACCCTG CCACGGTCCC TAAGCCGAAC CGCCGTAACG TAAGGATAAG
 P V D I D P A T V P K P N R R N V R I S>
 3180
 40 CGACGATCCT CAGACGGTGG TTGCTCGTCG GCGTCGGGAA AGGATCAGCG AGAAGATCCG
 D D P Q T V V A R R R R E R I S E K I R>
 3240
 45 AATTCTCAAG AGGATCGTGC CTGGTGGTGC GAAGATGGAC ACAGCTTCCA TGCTCGACGA
 I L K R I V P G G A K M D T A S M L D E>
 3300
 50 AGCCATACGT TACACCAAGT TCTTGAAACG GCAGGTGAGG ATTCTTCAGC CTCACTCTCA
 A I R Y T K F L K R Q V R I L Q P H S Q>
 3360
 55 GATTGGAGCT CCTATGGCTA ACCCTCTTA CCTTTGTTAT TACCACAACT CCCAACCTG
 I G A P M A N P S Y L C Y Y H N S Q P *>
 3420
 60 ATGAACTACA CAGAAGCTCG CTAGCTAGAC ATTTGGTGTC ATCCTCTCAA CCTTTTTCAT
 3480
 GTTGATATAT TATATATAGA TGCATAAAGA TTCGATCCAA GATTGTATGG GTGTTTAAAT
 3540
 ATTATTATTC TAAGATATAT GATGTACAAT TGTGTACCAA GTTTCTTTAT CTTGATATCA
 3600
 TATGCATAAA TAATTGGTGA ATAAAAAGAA GATATTGATT GTAAACAAAA AAAAGAAGAT
 3660
 ATTGATTGTT AATTAGGGTT TGATCATCTT GTATGAAAGC TTTGGCCTGC AAATTAATTT
 3720
 65 TCGATATATA TATATATATA TGGAGAATAT ATATCAAATA CTTTTTTAAT TTGACTATAA
 3780
 TTTGTATCAA TTATCTGAAT CTGATGAGTG TAGGTTATAT ATGGATTAGC AAAAAAGAAA

3840
ACAACCATTA TTACGCACCT ACATTAAAAA TCATCCACCA AAGAAGAAAC CATCCTCAAG
5 AGGGTTCCCT CTAGAG

SEQ ID NO:2 IND1 protein

10 M E N G M Y K K K G V C D S C V S S K
S R S N H S P K R S M M E P Q P H H L L
M D W N K A N D L L T Q E H A A F L N D
15 P H H L M L D P P P E T L I H L D E D E
E Y D E D M D A M K E M Q Y M I A V M Q
20 P V D I D P A T V P K P N R R N V R I S
D D P Q T V V A R R R R E R I S E K I R
I L K R I V P G G A K M D T A S M L D E
25 A I R Y T K F L K R Q V R I L Q P H S Q
I G A P M A N P S Y L C Y Y H N S Q P *

SEQ ID NO:3 IND1 5' promoter

30 60 CTCTAGACCA TCTACTATCC GGTTGTTGAC CCTTAAAGCT TTTGAAGACT ACTAGAATAA
120
TGCAAATACC ATATGTCCAT ATCCATCCTT TTCTTTTGTT TGAAGTGAAC ATTCTAATTT
35 180
TGTAAGAGAA AAAACCTTAT GTTAATATCA CCGTAGGCAA AAAAAATATC TCATCATATT
240
40 AAATTTTAT TATAAGATTA TACATTCTCT CGTTGTAAGA GTTACTCCAA TTGCAAGTGT
300
TGTATTAACT AATAAAAAGG ACGAAAGTAG GAAGCTTATA ATTAATTGAT GTTGCAATAGT
45 360
ACTGGTATAT TGTGATGAA TATAACAAGT ATGAACATTA ATGCATGAAA CGGGGTATTT
420
50 TGTCTGAAC TCATTAAAGG CAATGTGAAA AGAAGATGTG AGGTCTCATT TTGAAAATTT
480
ATCTTCTAGC TTTGTCGATT TTAAATCTAT GAAATGAACG CAACATATAG AAATTTTCATG
540
55 TGGACAACGA CATTAGACG GTATCTTAAT TAGACCGATT AATTAGTAAT ATACTTATAT
600
ATATAATTAG TGGTGATTAT AAGTTTACTT ATCCACTTGA GAATTTAAAC AATGGGCAAT
660
60 ACCTTAATGT CGAAGAAGC CGTCCCCACT TCGTGTAATG AGTTATGGGG GAGAGATCCT
720
GTAAATCGT CAAATAAAAC AACTTAAGAA CTAGAAATTG ACACCAAAAA TCATAAAGAG

	AACGTTGAAG	AAGTCATTTA	TCGTATCCAG	CTCATATTTT	CTAGCTAAGA	TCAAATCAAG	780
5							840
	GCCGTTGAAA	GGGCTTGTA	GAAAATGTCT	AAGAAACCGT	GGGGTTTAGA	AGAAAGACAA	900
10							960
	GAAATAGAAG	AACAATGATG	TTAAATTGCC	TATTTTGGTG	TATAGGAGTT	GTCAAAAGAG	1020
	GAGAGAGAGA	AGAAAATTAG	GTCAAAATAA	TGAGCACTAA	AAATGGAGAC	ATGTGTTGAG	1080
15							1140
	TAACTATTAC	AAGAGCGACT	TATGCTTCCT	TATGGCAATG	ATATCCAAAC	CAAAGTGCAA	1200
	CGCTCCTTTT	TTGCCCTAAT	TTGTAAGT	CTCTCTCCTT	CTTCGTCCTT	AGGAAAAACC	1260
20							1320
	CTAGAAATTT	AATCCCTTGT	TCTTGATCTT	GCTTTTGGAG	TAACCATGAT	TTTGACCACA	1380
	CACTATTTCT	TCTATCTTTT	GTGGTCTATA	GGATTTTGCT	TTATATGTGT	TTCTTGATT	1440
25							1500
	GCTCCGTACG	TACGTATACG	AATTTAAATG	GTTATAACAA	GGTTTATATA	AACTAGCACA	1560
	AATGAGTCCA	TGAAATTTGT	TAGCGAAAAA	GGTAGAAATA	TATTGAGTCT	TTAAACGGCA	1620
30							1680
	ATATATATAA	TTTTGCTGCA	AAACTTAGCT	TTAATCATGA	TCTAATGATA	TTTTCTTTAA	1740
35							1800
	TTTCCTTTGC	CAAATTAATC	ACATGCACGG	ATTTTGGCA	AGTTATGTGT	CGAATCTTTC	1860
	CATTCACACA	ACACTAAACT	TAATTAGAAC	TCTAGGAAAT	ATTTTAAAT	GACAACTTTA	1920
40							1980
	TCGAAAAAAA	TTTAGTTATG	AAAACAATTC	CAGAATTAAA	CATGAGCTAT	ATAATTTAAG	2040
	ATAAAATGAA	GTAATATTGA	TATGTATGTA	ATAACATATC	TGATTGCGGT	AAAAAAAAC	2100
45							2160
	ATATCTGATT	AAATGTTCA	TGCAGGCCCA	TGTCATATG	ATGTCATCAC	GTTTTTATTT	2220
50							2280
	TCACAATAAC	TAATATATAT	TCAAAAAAAT	AGTTTGTGCA	GATTAAATTT	TTTTTGGTGG	2340
	TCAGCTTTCT	CCAACCTACT	AAACTAGTTT	GGAATGTTCT	CTTCTTTATT	TTTCTTTTTC	2400
55							2460
	TTGATTTCTT	ATGTTTTTTA	TTTATGGAAT	TTTAAGACGG	ATTGTTTAGG	TCGTTTCTCT	2520
	CTTTTCTTGT	TTTCTAAAGT	TACTTTTGTA	AACTCATCTC	CTCCCAATTA	GACAGTCAAT	2580
60							2640
	CATATAGTTA	TCTTTTAATA	TATGTCTAGT	TGATAAAAAA	AATGAAAAAA	TACTGGTGGT	2700
65							2760
	AGTTCTACTA	ATGTTTGTGT	AAAAAATCTG	ATATTATGAA	TCTAATCAAT	TTCTTTGATC	2820
							2880

SEQ ID NO:5 SHP1 genomic

5 AGATCTGCAA CAGTGAAAAG AGAAAACAAA ATGGACTTGA AGAGGTTTTG ACAATGCCAG 60

AGATAATGCT TATTCCTTAA TATGTTGCCA GCCAAGTGTC AAATTGGCTT TTAAATATG 120

GATTCTGTGA TCAGTGGTCA TATTTGTGGA TCCAACGTAT TCATCATCAA GTTCTCAAGT 180

10 TTGCTTTCAG TGCAATTCTA ATTCACACGT TTAACCTTAA CATGCATGTC ATTATAATTA 240

CTTCTTCACT AAGACACAAT ACGGCAAACC TTCAGATTA TATTAATCTC CATAAATGAA 300

15 ATAATTAACC TCATAATCAA GATTCAATGT TTCTAAATAT ATATGGACAA AATTACACG 360

20 GAAGATTAGA TACGTATATT AGTAGATTTA GTCTTTCGTT TGTGCGATAA GATTAACCAC 420

CTCATAGATA GTAATATCAT TGTCAAATTC CTCTCGGTTT AGTCGCTAAA TTGTATCTTT 480

25 TTTAAGCCTA AAAGTAGTGT ATTCGCATAT GACTTATCGT CCTAACCTTT TTTTAAATTA 540

ACAAAAAAT CGAAAAGAAA ATAATCTGTT AAATATTTTT TAAGTACTCC ATTAAGTTTA 600

GTTTCTATTT AAAAAATGCT TGAAATTGGA CAGTTATGTT CAACAATTTT GAATCATGAG 660

35 CGATGTCTAG ATACTCAGAA TTTAATCAAG ATGTCTTATC AAATTTGTTG TCACTCGAGG 720

ACCCACGCAA AAGAAAAGAC TAATATGATT TTTATTTGGT CTGGATATTT TTGTAGAGGA 780

40 TGAAACTAAG AGAGTGAAAG ATTCGAAATC CACAATGTTC AAGAGAGCTC AAAGCAAAAA 840

GAAAAATGAA GATGAAGGAC TAAAGAACAA TAAGCAACTA CTTATACCCT ATTTCCATAA 900

45 AGGATTCAGG TACTAGGAGA AGTTGAGGCA AGTTNNNNNN NATTGATTCA AATTTTCATT 960

50 TATTTTTACA ATTTAATTCA CCTAAGTTAT TATGCATTTT TCATCATTGG TACATTTTCT 1020

GTATAGCGTA TTTACATATA TGAAATAAAT TAAATATGTC CTCACGTTGC AAGTAGTTAA 1080

55 TGAATGTCCC CACGCAAAAA AAAATCCCTC CAAATATGTC CACCTTTTCT TTTCTTTTAA 1140

ATTCCAAAAT TACCATAAAC TTTTGGTTTA CAAAAGATT CTAGAAATTG AGGAAGATAT 1200

60 CCTAAATGAT TCATGAATCC TTCAATAATC TGAAGTTTGC GATATTTTCG ATTTTCTTCA 1260

65 AGAGTTGCGA TATTTGTAAT TTGGTGACCT TAACTTTTTT TTGATAAAGA GTAAACGTTT 1320

5 AAAGCCTTCC CTTTTCAGGA AAAAAACAT TTCGAAATTA ACTCTAATTA ATCAAGAATT
 TCCTACAATG TATACATCTA ATGTTTTTTC CGCGATCTTA CTTATTAGTG TGAGGGGTAC 4140
 AATTGAAAGG TACAAGAAAG CTTGTTCCGA TGCCGTCAAC CCTCCTTCCG TCACCGAAGC 4200
 10 TAATACTCAG GTACCAATTT ATATTGTTG ATTCTCTTG TTTTATCTTC TTCTTTTCAT 4260
 TATATATATG ATCAACAAAA AATATAACCT ACAAAAAGAG AGAGTTCAAG GAAATGCATT 4320
 15 GAAACGGTTT CGTTATGGTG TTTGAATACA TGGATTTTTG AAGTACTATC AGCAAGAAGC 4380
 CTCTAAGCTT CGGAGGCAGA TTCGAGATAT TCAGAATTCA AATAGGTAAT TCATTAACCT 4440
 20 TTCATGAACT CTTCGATTG GTATTAGGTC ACTTAATTG GTGTCGGTCC AAAAGTCCGC 4500
 TTGTAGTTT CTTTAGAAGT TGTGTTGTT AATGTTTCATG TTTACAAAT GAAGGCATAT 4560
 TGTTGGGGAA TCACTTGGTT CCTTGAACCT CAAGGAACTC AAAACCTAG AAGGACGTCT 4620
 30 TGAAAAAGGA ATCAGCCGTG TCCGCTCCAA AAAGGTAAAA TCTACGTTGC TCTCTCTCTG 4680
 TGTCTCTGTC TCTCTCTCTA TATATAGTCC CTTAGTTTAT ATAGTTCATC ACCCTTTTGT 4740
 35 GAGAATTTTG CAGAATGAGC TGTTAGTGGC AGAGATAGAG TATATGCAGA AGAGGGTAAG 4800
 AACGTTTCTC CCATTCCAAG TAATTAGATC TTTCTTCGTC TTTGTGAGGG TTTGAGTTT 4860
 CCCATAAATC ATGTGTAGGA AATGGAGTTG CAACACAATA ACATGTACCT GCGAGCAAAG 4920
 45 GTTAGCCACG TTCTGTTCCA AATCTTAATC TCAATATCTA CTCTTTTCTT CATGTATATA 4980
 CTAAGATAAC GTGAATAACA AGAAAACCTT TGTGTTTGGG TTTAATAGAT AGCCGAAGGC 5040
 50 GCCAGATTGA ATCCGGACCA GCAGGAATCG AGTGTGATAC AAGGGACGAC AGTTTACGAA 5100
 TCCGGTGTAT CTTCTCATGA CCAGTCGCAG CATTATAATC GGAATATAT TCCGGTGAAC 5160
 CTTCTTGAAC CGAATCAGCA ATTCTCCGGC CAAGACCAAC CTCCTCTTCA ACTTGTGTAA 5220
 60 CTCAAACAT GATAACTTGT TTCTTCCCT CATAACGATT AAGAGAGAGA CGAGAGAGTT 5280
 CATTTTATAT TTATAACGCG ACTGTGTATT CATAGTTTAG GTTCTAATAA TGATAATAAC 5340
 65 AAAACTGTTG TTTCTTTGCT TAATTACATC AACATTTAAA TCCAAAGTTC TAAAACACGT 5400

5460
CGAGATCCAA AGTTTGTCAT ACAAGATTAG ACGCATACAC GATCAGTTAA TAGATTTTAA
5520
5
GTGCCTTTTA ATATTTACAT ATAGTTGCAG CTTGATTAG ATCATGTCCA CCAAACACTC
5580
10
ACAATTAGAG ACAAGCAAAA CTATAAACAT TGATCATAAA ATGATTACAA CATGTCCATA
AATTAATTAT GGATTACAAA AATAAAACT TACAAAAGAT CT

SEQ ID NO:6 SHP2 genomic

60
15
GAATTCGTAA CAGAATTTAG TGAATAATAT TGTAATTACC AGGCAAGGAC TCTCCAAACG
120
GATAGCTCGA ATATCGTTAT TAAAGAGTAA ATGATCCAAT ATGTAAGCCA TTGTTGATCA
180
20
TCTAACATTG TTGGACTCTC TATTGCTCGA AATGATGCAT ACCTAATCAT TTATTCAGTT
240
25
AACTATCAAG TTGCATTTGT AAAAACCAAA CATTTAAATT CAGATTTGAT ATCACTTACA
300
GAGGATAGAG AAGCATGACT CCAGGCCTGC ATGCAACAAG AAAAAGGAAG AAAATAATGT
360
30
TAAAAATTG ACAAATATAG TGTTTATTTT TATTATATGA GACAGAATTT GAATAAAATC
420
CTACCCAACT AGAGCATCAA AACGTTTGC AATCGCAATA ATGAAACCCA TTTTCTTTTT
480
35
GAGTTTTTAC TCTCTTTCA ACAGAACTT TCTCAAACGT CTTTAGCACT GTGACGTTAG
540
40
ATATATACAC AAAAGCTTGA AATTTCTTCA AGCAAAAGAA TCTTTGTGGG AGTTAAGGCA
600
ACAAGCCAGG TAAAGAATCT CCAACGCATT GTTACGTTTT CATGAACCTA TTTATTATAT
660
45
GTTCTAAGAA AGAAAAAAT ATCTCAAAGT AAACGTTGGA AATTTTCTGA TGAAGGGAAA
720
TCCAAAGTCT TGGGTTTAGT ATCCCTATGA ATGGTATTTG GAATATGTTT TCGTCAAAAC
780
50
AAAAGATTCT TTTCTTTTTC ACAAGAGTTA GTGATCAATA ACTTATGCAC TAATTAATGA
840
55
GATTGGACGT ATACACAATT TGATTATGAT ACTTGAGTAA AAATCACCTG TCCTTTAATT
900
TGGAATCTC TCTTTCTTAC CCATTTATAT ACTACTTCTT TTCATTAAAA TTAAATTTCA
960
60
ATTATCAATC ATCGTTCAAT TTGATAAAGA TTAAACATTT TTTGTCACAG GGCTAGTAAA
1020
AGCAATCTTT ACATAATTCA TCTTTCTTAC ATATATATAT TACCTTTTTC TTCATTAGTA
1080
65
TTCTATTTGA TTATGATTAT TTTGTCATAA AGCTAGTAAA TTAAACACTC GATATGAGAA

1140
TTATATTACT TCACGCTAAT TAACTCTTAA CACAACAAGA ACTAGTGCAT ATTCAACTTT

5 1200
CAAAGCATAT ACTATATATT GAGAATATAG ACCACGAAAG TCAATCAAAA GACCTACCAG

10 1260
CTCTCATCAA GTTCTTTCTT GAAATGATTT TGCAGAATTT CCAACTTAAT TAATTGACA

15 1320
TGAATGTGAA AATGTGTGTT GCTCGTTAAG AAAATTGAAT AGAAGTACAA TGAAAATGAT

20 1380
GAGGAATGGG CAAAACACAA AAGAGTTTCC TTTCGTAACT ACAATTAATT AATGCAAATC

25 1440
TGAGAAAGGG TTCATGGATA ATGACTACAC ACATGATTAG TCATCCCCG TGGGCTCTCT

30 1500
GCTTTCATTT ACTTTATTAG TTTCATCTTC TCTAATTATA TTGTCGCATA TATGATGCAG

35 1560
TTCTTTTGTC TAAATTACGT AATATGATGT AATTAATTAT CAAAATAAAT ATTCAAATTG

40 1620
CCGTTGGACT AACCTAATGT CCAAGATTAA GACTTGAACA TAAGAATTTT GGAAAACTA

45 1680
AACCAGTTAT AATATATACT CTAAATTGC CATTTCTGAA CACAACCAAA TAATAATATA

50 1740
TACTATTTAC AGTTTTTTTT AATTGGCAAG AACACTGAAA TCTTATTCAT TGTCTCGCTT

55 1800
GGTAGTTGAC AAGTTATAAC ACTCATATTC ATATAACCCC ATTCTAACGT TGACGACGAA

60 1860
CACTCATATA AACCACCCAA ATTCTTAGCA TATTAGCTAA ATATTGGTTT AATTGGAAAT

65 1920
ATTTTTTTTA TATATAAAAT GCCAGGTAAA TATTAACGAC ATGCAATGTA TATAGGAGTA

2040
GGGCAATAAA AAGAAAAGGA GAATAAAAAG GGATTACCAA AAAAGGAAAG TTTCCAAAAG

2100
GTGATTCTGA TGAGAAACAG AGCCCATACC TCTCTTTTTT CCTCTAAACA TGAAAGAAAA

2160
ATTGGATGGT CCTCCTTCAA TGCTCTCTCC CCACCCAATC CAAACCCAAC TGTCTTCTTT

2220
CTTTCTTTTT TCTTCTTCT AATTGATAT TTTCTACCAC TTAATTCCAA TCAATTTCAA

2280
ATTTCAATCT AAATGTATGC ATATAGAATT TAATTAAG AATTAGGTGT GTGATATTTG

2340
AGAAAATGTT AGAAGTAATG GTCCATGTTC TTTCTTTCTT TTTCTTCTA TAACACTTCA

2400
GTTTGAAAAA AACTACCAA ACCTTCTGTT TTCTGCAAAT GGGTTTTTAA ATACTTCCAA

2460
AGAAATATTC CTCTAAAAGA AATTATAAAC CAAAACAGAA ACCAAAAACA AAAAATAAAG

3840
 TAATCTATAA GTCAAGTCTC ATATCATGGA TCTAAGTTAA AACTAGTAAA TTTGTAGTTT
 3900
 5 CAATGTGAAC TTTCAACAAC ACTAAAGAAC TGATCTGAAG TTTATAATGG ACATGACTAA
 3960
 TTTGATTAAAC AAAAGAGGAA TGCATTATGT ATGTAGAAAC ATGTGATATA TATATGTTTC
 4020
 10 TATTATCAAA AGTGTAGTTA ACTTTCTTAT TTCAAACACC CTCATGCTTT AGTAGTATCT
 4080
 15 TACTTTTGAC ATTTCTCAAC TTCAGCTTTC CATTATACAA CAGCACAATG TAAATTACTT
 4140
 GTATATGAAT ATGAAAGCAT AACGTTATGC AAAGATTCTT AGCTTTTCTT TTTCTGTTTT
 4200
 20 GCAAAAGATT TACAAATATC ATGTTCTTGG TAAAAACATA CTTGCCTCAG CCACATATGC
 4260
 ATGTAAATGT AATGTTCAAA TATTAATTCA GGAAAAACAA AGAAGAAGCA AAATTAGCTT
 4320
 25 CTAGAGTAGG GAATCTATTG ACTTGACCTG AAAATCACTT CTTTTTCTTA AAGCCTAGTA
 4380
 30 GTGAATTTTT TAATCTAATT AGGCCAAAAT ATATACTAGC CTAAATATA ATTTGGATT
 4440
 TGTGTCGTAC ATAAATTGGG ACCAATTCCA ATTAATAAG AGCATATGCA ATTCAAATTC
 4500
 35 TTTTTATTTT CTTCTCCGAT TTGCTACTTC TTTCTTTTGT ATGTTTTCAA ATTAGGATTA
 4560
 CACTTTTTTG GGGAAGTACA CATTAGGGTC TTCTCGAACT TTGATTATAC ATATATATAT
 4620
 40 ATATATATAT ATATAACTTT GTGAGATGTC ACTGTTAATA GATAATAGGC AATAACAATA
 4680
 ATATCCAAAA AAGAAGGCGC AAACAAATCA TATACTATAT GGTACTGGTC CATTCACTAT
 4740
 45 TTTGTCGGTT GAATTTAAGG TTTGGCGTAC AACTTTGTT TCAAACCTT ATTATCCGT
 4800
 50 CTTTCTGTGT GTTTTGATA TCCAGAAGAT AAAAATATCA ATTTCTTTAA CGACTTCATA
 4860
 TATATATATA TATATATATA TATATATATT TTTCTCTTCT GGTTTTAGTG TTTGAATCCA
 4920
 55 ACAGTTATAG TTTCGTGTGT CTTTGTTTTA CTTGTGGTGG TTTAAGTTG AGATTTTCAC
 4980
 60 CGATTGCATC TATTACATA TATAGCTACC ACAAAAAAGA TTGCATTTA AAATCTTTTC
 5040
 CTTTGTGTGA ATGTTGATGA AGTGTGAGAG GAACAATAGA AAGGTACAAG AAAGCTTGCT
 5100
 65 CCGACGCCGT TAACCCCTCCG ACCATCACCG AAGCTAATAC TCAGGTTAGC TTTTAATTAA
 5160
 TACACCTAGC TAGCTAGTTC GTTAATTACT TAATTTCTTC TTCTTTTAGT TATCTGACCT

5
10
15
20
25
30
35
40
45
50

TTTTTTCACC TCTTGTAACA ATGATGGGAT CGAAATTGAT GAAGTACTAT CAGCAAGAGG 5220
CGTCTAAACT CCGGAGACAG ATTCGGGACA TTCAGAATTT GAACAGACAC ATTCTTGGTG 5280
AATCTCTTGG TTCCTTGAAC TTTAAGGAAC TCAAGAACCT TGAAAGTAGG CTTGAGAAAG 5340
GAATCAGTCG TGTCCGATCC AAGAAGGTAC ATCTACTAAT CTCCATCAAT CTCCTTATCA 5400
TTGAATATAT ATCCATCTGA TTCTTGCCCG TTATATTTGG TTTTCTCTC CAGCACGAGA 5460
TGTTAGTTGC AGAGATTGAA TACATGCAA AAAGGGTAAA AGTAAAACCT ATCTTCCTTC 5520
ACAATGAACCT ACCCCTACTT TATTAGCAAC TTCTCTTTCT GATGATCATC TTTTTTATTT 5580
TCTGTTGTCG CTTGCATTGT AGGAAATCGA GCTGCAAAAC GATAACATGT ATCTCCGCTC 5640
CAAGGTTTTA TACATAACTC TTTTGGCAT TTTTGATCAT CATTTTTTTC CGGTAGACAA 5700
TCTCTTGATG TGCAAATTCT AAATATCTCT GCAGATTACT GAAAGAACAG GTCTACAGCA 5760
ACAAGAATCG AGTGTGATAC ATCAAGGGAC AGTTTACGAG TCGGGTGTTA CTTCTTCTCA 5820
CCAGTCGGGG CAGTATAACC GGAATTATAT TGCGGTTAAC CTTCTTGAAC CGAATCAGAA 5880
TTCCTCCAAC CAAGACCAAC CACCTCTGCA ACTTGTTTGA TTCAGTCTAA CATAAGCTTC 5940
TTTCCTCAGC CTGAGATCGA TCTATAGTGT CACCTAAATG CGGCCGCGTC CCTCAACATC 6000
TAGTCGCAAG CTGAGGGGAA CCACTAGTGT CATAACGAACC TCCAAGAGAC GGTTACACAA 6060
ACGGGTACAT TGTGATGTC ATGTATGACA ATCGCCCAAG TAAGTATCCA GCTGTGTTCA 6120
GAACGTACGT CCGAATTC